Singular Value Decomposition (SVD)

L. Olson

Department of Computer Science
University of Illinois at Urbana-Champaign

objectives

- Construct a *singular value decomposition* or SVD
- Look at some problems the singular values are useful
- Highlight several properties of the SVD
- What do the singular values mean?
- How do then impact our numerics?
- What is the cost of computing them?

svd: motivation

SVD uses in practice:

1. Search Technology: find closely related documents or images in a database
2. Clustering: aggregate documents or images into similar groups
3. Compression: efficient image storage
4. Principal axis: find the main axis of a solid (engineering/graphics)
5. Summaries: Given a textual document, ascertain the most representative tags
6. Graphs: partition graphs into subgraphs (graphics, analysis)

svd: singular value decomposition

SVD takes an $m \times n$ matrix A and factors it:

$$A = USV^T$$

where U ($m \times m$) and V ($n \times n$) are orthogonal and S ($m \times n$) is diagonal.

Definition

A is orthogonal if $A^TA = AA^T = I$.

S is made up of "singular values":

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r \geqslant \sigma_{r+1} = \cdots = \sigma_p = 0$$

Here, $r = rank(A)$ and $p = min(m, n)$.

we want...

$$A = \begin{bmatrix} \vdots & \vdots & \vdots \\ u_1 & \dots & u_m \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix} \begin{bmatrix} \dots & v_1^T & \dots \\ \dots & \vdots & \dots \\ \dots & v_n^T & \dots \end{bmatrix}$$

diagonalizing a matrix

We want to factorize A into U, S, and V^T. First step: find V. Consider

$$A = USV^T$$

and multiply by A^T

$$A^TA = (USV^T)^T(USV^T) = VS^TU^TUSV^T$$

Since U is orthogonal

$$A^TA = VS^2V^T$$

This is called a similarity transformation.

Definition

Matrices A and B are similar if there is an invertible matrix Q such that

$$Q^{-1}AQ = B$$

Theorem

Similar matrices have the same eigenvalues.

proof

$$\begin{aligned} Bv &= \lambda v \\ Q^{-1}AQv &= \lambda v \\ AQv &= \lambda Qv \\ Aw &= \lambda w. \end{aligned}$$

Further, if v is an eigenvector of B, Qv is an eigenvector of A.

so far...

Need $A = USV^T$

Look for V such that $A^TA = VS^2V^T$. Here S^2 is diagonal.

If A^TA and S^2 are similar, then they have the same eigenvalues. So the diagonal matrix S^2 is just the eigenvalues of A^TA and V is the matrix of eigenvectors. To see the latter, note that since S^2 is

diagonal, the eigenvectors are e_i, and V^Te_i is just the ith column of V^T.

similarly...

Now consider

$$A = USV^T$$

and multiply by A^T from the right

$$AA^T = (USV^T)(USV^T)^T = USV^T VS^T U^T$$

Since V is orthogonal

$$AA^T = US^2 U^T$$

Now U is the matrix of eigenvectors of AA^T.

in the end...

We get

$$A = \begin{bmatrix} \vdots & \vdots & \vdots \\ u_1 & \dots & u_m \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix} \begin{bmatrix} \dots & v_1^T & \dots \\ \dots & \vdots & \dots \\ \dots & v_n^T & \dots \end{bmatrix}$$

example

Decompose

$$A = \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix}$$

First construct $A^T A$:

$$A^T A = \begin{bmatrix} 2 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}$$

Eigenvalues: $\lambda_1 = 8$ and $\lambda_2 = 2$. So

$$S^2 = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \Rightarrow S = \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$$

example

Now find V^T and U. The columns of V^T are the eigenvectors of A^TA.

- $\lambda_1 = 8$: $(A^TA - \lambda_1 I)v_1 = 0$

$$\Rightarrow \begin{bmatrix} -3 & -3 \\ -3 & -3 \end{bmatrix} v_1 = 0 \Rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} v_1 = 0 \Rightarrow v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$$

- $\lambda_2 = 2$: $(A^TA - \lambda_2 I)v_2 = 0$

$$\Rightarrow \begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix} v_2 = 0 \Rightarrow \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} v_2 = 0 \Rightarrow v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$$

- Finally:

$$V = \begin{bmatrix} -\sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$

example

Now find U. The columns of U are the eigenvectors of AA^T.

- $\lambda_1 = 8$: $(AA^T - \lambda_1 I)u_1 = 0$

$$\Rightarrow \begin{bmatrix} 0 & 0 \\ 0 & -6 \end{bmatrix} u_1 = 0 \Rightarrow \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} u_1 = 0 \Rightarrow u_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

- $\lambda_2 = 2$: $(AA^T - \lambda_2 I)u_2 = 0$

$$\Rightarrow \begin{bmatrix} 6 & 0 \\ 0 & 0 \end{bmatrix} u_2 = 0 \Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} u_2 = 0 \Rightarrow u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- Finally:

$$U = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

- Together:

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} -\sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$

svd: who cares?

How can we actually *use* $A = USV^T$? We can use this to represent A with far fewer entries...

Notice what $A = USV^T$ looks like:

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T + 0 u_{r+1} v_{r+1}^T + \cdots + 0 u_p v_p^T$$

This is easily truncated to

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$

What are the savings?

- A takes $m \times n$ storage
- using k terms of U and V takes $k(1 + m + n)$ storage